Comparative study of texture features in OCT images at different scales for human breast tissue classification

Yu Gan, Xinwen Yao, Ernest Chang, Syed Bin Amir, Hanina Hibshoosh
Sheldon Feldman, Christine P. Hendon

Abstract— **Breast cancer is the second leading cause of death in women in the United States due to cancer. Early detection of breast cancerous regions will aid the diagnosis, staging, and treatment of breast cancer. Optical coherence tomography (OCT), a non-invasive imaging modality with high resolution, has been widely used to visualize various tissue types within the human breast and has demonstrated great potential for assessing tumor margins. Imaging large resected samples with a fast imaging speed can be accomplished by under-sampling in the spatial domain, resulting in a large image scale. However, it is unclear whether there is an impact on the ability to classify tissue types based on the selected imaging scale. Our objective is to evaluate how the scale at which the images are acquired impacts texture features and the accuracy of an automated classification algorithm. To this end, we present a comparative study of texture features in OCT images at two image scales for human breast tissue classification. Texture features and attenuation coefficients were inputs to a statistical classification model, relevance vector machine. The automated classification results from the two image scales were compared. We found that more informative tissue features are preserved in small image scale and accordingly, small image scale leads to more accurate tissue type classification.**

I. INTRODUCTION

Breast cancer is the second leading cause of death in women in the United States due to cancer in 2015 [1]. There is an unmet need to detect breast cancer especially at the early stage, where interventions have the potential to be minimally invasive. Due to its importance, many imaging modalities have been utilized to image breast tissue for cancer detection. The most established technique for breast cancer detection is the mammogram [2], which provides tomographic images of the breast by using low-dose X-rays. Automated whole breast ultrasound (AWBU) [3] has been suggested as a complementary modality to mammograms for breast cancer detection in dense-breasted women. Diffuse optical tomography (DOT) [4] has the ability to detect suspicious lesion in breast. Overall, mammogram, AWBU, and DOT are non-invasive and are good at screening large tumor masses.

Regarding more detailed morphological information, Optical coherence tomography (OCT), a non-invasive modality providing resolution at micron-meter level, has been investigated as a promising imaging modality to detect
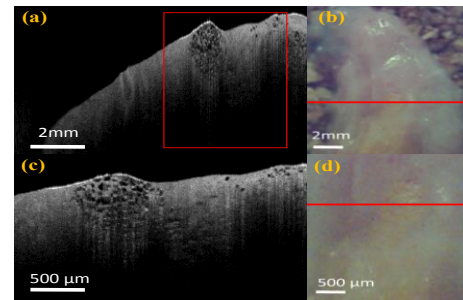
Figure. 1 OCT image at different scale. (a) OCT B-scan at 10 mm field of view; (b) white light image at 10 mm field of view. 1 (c)OCT B-scan at 4 mm field of view; (d) OCT image at en face plane at 4 mm field of view. Red box in (a) corresponds the whole region imaged at (c). Red line in (b) and (d) correspond the plane where we took the B-scans in (a) and (c).

cancerous region within breast tissue, either through needle based probes [5] or imaging resected tissue samples [6-10]. Features within normal tissue, benign lesions, and malignant lesions were captured by an ultrahigh resolution OCT and showed strong histological correlations [6] to aid manual diagnosis. Functional extensions of OCT, such as optical coherence elastography, have been used to identify the mechanical property [7] of tumor. Automated algorithms [8-10] have been developed to classify tumor region with 70-90 % accuracy. There is a potential for the use of OCT to guide surgical margin assessment during clinical procedure.

Given the fact that the resected tissues are larger than the standard OCT field of view (4mm), there is a need to use OCT to image large areas within a short time period, resulting in a larger image scale. We aim to evaluate how the scale at which the images are analyzed impacts texture extraction and classification results. In this paper, we propose a comparative study on the automated classification of human breast tissue features at different image scales. We imaged human breast tissue at two image scales and extracted texture features for each scale. Texture features were input to a statistical classification model, relevance vector machine (RVM), and the classification results from two scales were compared.

II. METHODOLOGY

A. Data collection

Three-dimensional OCT volumetric images were taken on samples from 19 patients at Columbia University Medical Center, including both healthy breast tissue from breast reductions and from mastectomies. Samples were acquired from the Department of Pathology and Cell Biology's tissue bank within 12 hours after resection and stored in phosphate buffered saline until imaging. After imaging, samples were fixed in 10% formalin for ~24 hours and then in ethanol (20%) for ~24 h for H&E histology.
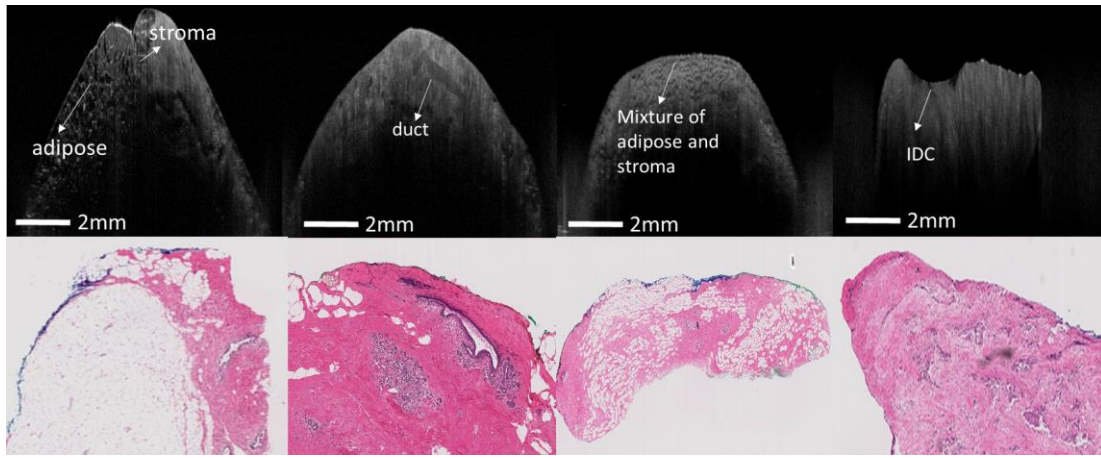
Figure 2. Typical breast tissue types: (a) adipose tissue and stroma in OCT images; (b) duct structure in OCT image; (c) mixture structure of adipose and stroma in OCT image; (d) invasive ductal cacinoma (IDC) structure in OCT images. (e)-(h) histological images of the same regions in (a)-(d).

All samples were imaged *ex vivo* at room temperature, using a commercial OCT system, Telesto I (Thorlabs GmbH, Germany). The light source of the system is centered at 1325 nm, and the axial and lateral resolutions of the system are 6.5 μm and 15 μm in air respectively. All data were acquired at a 28 kHz line rate. In our experiments, we have a subset of the samples imaged with two scales. For large-scale images, each B-scan was 2.52 mm × 10 mm laterally corresponding to 512 pixels × 600 pixels, and for small-scale, that was 2.52 mm × 4 mm laterally corresponding to 512 pixels × 600 pixels as shown in Fig. 1(a) and Fig. 1(c). Ten millimeters was chosen to try and obtain a full view of the sample while staying within the lateral range of the system's sample arm. The dataset acquired with the large scale imaging setting under-sampled the spot size of the OCT objective.

B. Feature extraction

With the guidance from histological analysis, we identified five tissue types from the OCT images we acquired, including adipose tissue, stroma, duct, mixture of adipose and stroma, and invasive ductal carcinoma (IDC). As shown in Fig. 2(a, e), fibrous stroma appears highly scattering and adipose tissue shows a honeycomb pattern. Duct structure, shown in Fig. 2(b, f), has lower scattering than the surrounded stroma. A mixture of adipose tissue and stroma, as seen in Fig. 2(c, g), has higher scattering than adipose and shows more heterogeneous than normal fibrous stroma. In Fig. 2(d, h), an example of IDC is shown, where the image is heterogeneous due cancer cells infiltrating into the breast tissue.

In our analysis, we used raw OCT images extracted from the Thorlabs system. The four consecutive B-scans were averaged to reduce the speckle noise. In addition, we performed histogram normalization on all OCT images. Then, we specified three sub-regions in each B-scan for the two image scales and identified the tissue types based on histological analysis. For each region, we extracted texture features, including entropy, local standard deviation, homogeneity from grey level co-occurrence matrix (GLCM) [11], and coarseness from texture feature coding number (TFCN) [12]. We drew the histogram of the OCT intensity in each region and calculated the ***entropy*** within each region based on the distribution obtained from histogram. We also studied the ***local standard deviation*** within an area of 7 pixels

× 7 pixels and computed the mean of local standard deviation within the whole sub-region. In addition to the histogram of intensity, we considered the relative position of pixels in each region and investigated the feature in GLCM. In GLCM, the relative position of pixels was used to construct a co-occurrence matrix. Setting a level of 16, we obtained a probability showing the occurrence of paired intensity. The ***homogeneity*** of the sub-region was calculated by measuring the spatial closeness of distribution of entries in the diagonal of co-occurrence matrix. Moreover, we analyzed the texture using TFCN, in which we encoded the change of direction trend over its surrounding 8 pixels. We further drew the histogram of the texture feature coding numbers. The feature ***coarseness*** was extracted by calculating the ratio of number of pixels having the maximum variations over 8-connectivity neighborhood over the total number of pixels within the region. In addition to texture features, we extracted the attenuation coefficients (mm^{-1}) from averaged A-lines, which is measured based on the method mentioned in [13]. In total, we had five features for each region.

C. Classification algorithm

For each unknown region, the extracted feature formed a feature vector. We then input the features to an RVM [14, 15] to classify tissue types. Since the structure of breast tissue is very heterogeneous, we are interested in not only the specific tissue that the region belongs but also other possible tissue types within the same region. Therefore we choose a probabilistic model RVM. Compared with support vector machine, RVM has the advantages of requiring less number of vectors and loosen condition for kernel function. For each feature vector x, we determined its probability to fall into a specific tissue type c by the following equation:

$$p(c=1 \mid w, x) = \sigma(\sum_{i}^{B} \omega_i \phi_i(x)) \qquad (1)$$

where $\phi(x)$ is a kernel function, w is the corresponding weight of each kernel function, $\sigma(\bullet)$ is a sigmoid function, and B is the number of vectors. Here, we use a squared exponential kernel with parameter $\lambda = 10^{-3}$. To determine the weight of each function, we used a zero mean Gaussian prior as

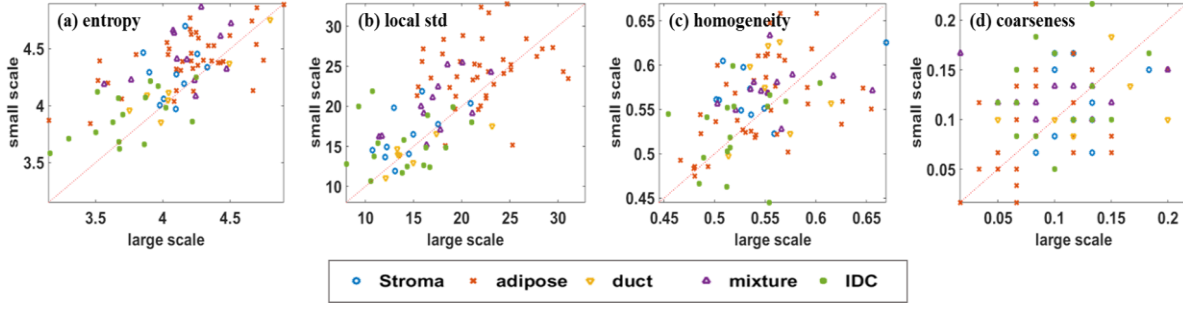$$p(w_i \mid \alpha_i) = N(w_i \mid 0, \alpha_i^{-1}) \qquad (2)$$

Figure. 3 Feature comparison at two image scale (large-scale: 2.52 mm × 10 mm; small-scale: 2.52 mm × 4 mm) for the features of (a) entropy (b) local standard deviation (c) homogeneity and (d) coarseness. More than 70% of points locate at the upper region of diagonal.
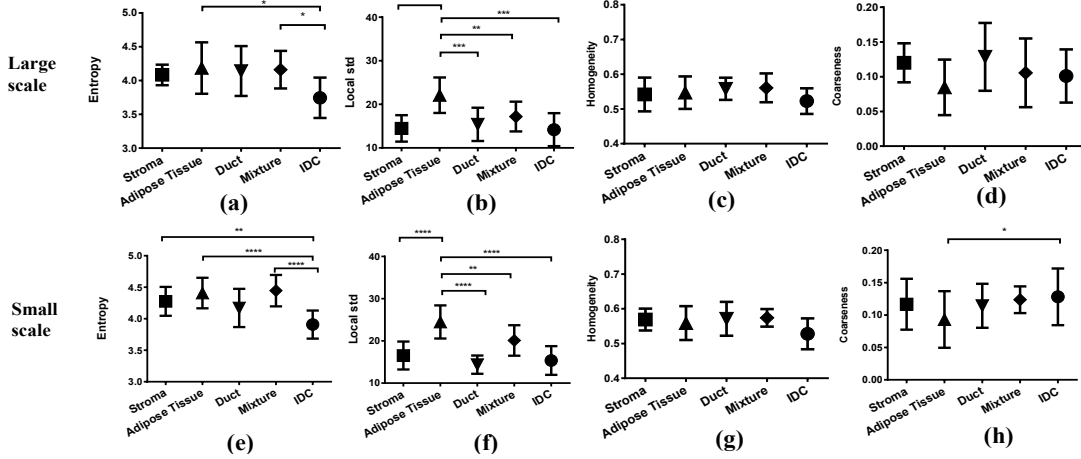

Figure. 4 Statistical comparison of texture features in entropy (a, e), local standard deviation (b, f), homogeneity (c, g), and coarseness (d, h). Fig (a) to (d) are for the large-scale while fig (e) to (h) are for the small-scale. In general, the features show higher significance level in small-scale images (*: p<0.05; **: p< 0.001; ***: p<0.0001; ****: p<0.00001).

Here, we used a hyper-parameters α_i to determine the distribution. For any dataset with known tissue types, $D = \{(x^n, c^n), n = 1,2,..., N\}$, we set α_i by maximizing the marginal likelihood

$$p(D \mid \alpha) = \int \prod_{n=1}^{N} p(c^n \mid x^n, w)(\frac{\alpha}{2\pi})^{B/2} \exp(-\frac{\alpha}{2} w^T w)dw \quad (3)$$

In this paper, we use Gull-MacKay method to update α_i.

To update w, we calculated the derivative of the expectation E[w] and the Hessian matrix of the weights H. In particular, we used a Newton update method to estimate the weight as following:

$$W^{NEW} = w - H^{-1}(\nabla E) \quad (4)$$

Hyper-parameters and weights were updated alternatively until we obtained a converged w. Following the training, we estimated the probability of each unknown region belonging to specific tissue type.

III. EXPERIMENTS AND RESULTS

A. Texture comparison among different scales

As mentioned in Section II.A, we extracted features in OCT images from two scales and compared the extracted features. Fig. 3 is a plot of the feature values from the large-scale images versus that from the small-scale. If the features obtained from the large-scale image match the features obtained from the small-scale, all points should be aligned along the diagonal in all of the plots. However, we found large variations from the diagonal. We fitted the spots in each plot and found the R^2 values of matching are between 0.01 and 0.86. The highest R^2 was obtained when matching the entropy features in duct structure while the lowest R^2 was obtained when matching TFCN features in mixture of adipose and stroma structure. Moreover, tissue types showed great differences with respect to the scales. We found that more than 70% of the points locates on the upper region of diagonal, indicating that for the features of entropy and homogeneity, the values extracted from small-scale images are larger than those from the large-scale. In general, there are more texture information in small-scale than large-scale images. For attenuation coefficient, we found that adipose tissue has the highest value and stroma has the lowest value.

The impact of image scale varied among tissue types. For example, in Fig. 3 (b), the duct tissue has a smaller variation than IDC tissue. Interestingly, the stroma tissues are along the diagonal in local standard deviation but are well beyond the diagonal in homogeneity. This means extracted homogeneity is more sensitive to the image scale than local standard deviation. Similarly, with a large variation of entropy value, features showed in small-scale images may not be observable in large-scale images.

B. Statistics of textures among tissue types

We further compared the statistics of each feature in Fig. 4. The values of each feature were plotted as mean with standard
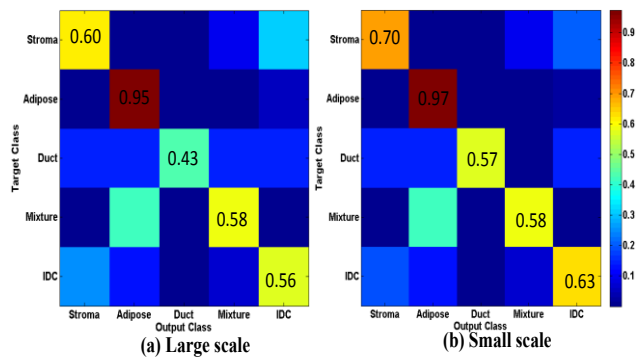
Figure 5. Confusing matrix obtained from RVM classification test from (a) large-scale image dataset; (b) small-scale image dataset. The overall accuracy from small-scale images was higher than that from large-scale.

deviation. One-way analysis of variance (ANOVA) with Tukey multiple comparison test were performed to examine the differences among the tissue types for each of the extracted features. We found that the texture value varies among the five tissue types as shown in Fig. 3. Generally, we observed greater significant differences in features between tissues types measured from the small-scale images. For example, in coarseness, the adipose tissue is not significantly different from IDC in large-scale image, but statistically they are different in small-scale. Moreover, for the tissue types that are different, the level of significance increases in small-scale images, as indicated by smaller p values, if comparisons are made between Fig. 4(a) and (e) or between Fig 4. (b) and (f).

C. Classification results

We incorporated the extracted features into the classification model we used in [14] and verified the classification results from 84 regions from 28 B-scans with both image scales following a leave-one-out test. The results were shown in Fig. 5. We found the overall accuracy from small-scale images (78.6%) was higher than that from large-scale images (73.8%). Adipose tissue is very distinctive and the classification results are not sensitive to image scale. However, the detection accuracy of other tissue types, such as duct and IDC, is improved by using small-scale imaging because small-scale imaging can provide more detailed morphological information and feature information.

IV. DISCUSSION AND CONCLUSION

We presented a comparative study of extracted tissue features at different scales of human tissue classification. We compare the correlation between extracted tissue features in two image scales and statistically study the difference among tissue types regarding the extracted features. In addition, we utilized a probabilistic model (RVM) to evaluate the classification performance of human breast tissue classification. We found that in general more tissue features are preserved in small imaging scale and accordingly, small imaging scale leads to more accurate tissue classification.

Due limited sample size, we did not have enough samples to analyze the impact of image scale on ductal carcinoma in situ (DCIS). In the future, we plan to extend our study into identification of both DCIS and IDC and the tumor stage of each cancerous type. In addition, we will evaluate the impact of image scales on classification with data acquired using a high-resolution OCT system at 800 nm [16, 17]. With improvement in processing speed for real time analysis, automated analysis of resected samples or catheter-based interrogation of human breast tissue with optical coherence tomography shows promise for identifying important features related to breast cancer.

REFERENCE

[1] A. C. Society, "Cancer Facts & Figures 2015. Atlanta: American Cancer Society," 2015.

[2] L. Nyström, S. Wall, L. E. Rutqvist, A. Lindgren, M. Lindqvist, S. Rydén, et al., "Breast cancer screening with mammography: overview of Swedish randomised trials," The Lancet, vol. 341, pp. 973-978, 4/17/ 1993.

[3] K. M. Kelly, J. Dean, W. S. Comulada, and S.-J. Lee, "Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts," European radiology, vol. 20, pp. 734-742, 2010.

[4] M. L. Flexman, H. K. Kim, J. E. Gunther, E. A. Lim, M. C. Alvarez, E. Desperito, et al., "Optical biomarkers for breast cancer derived from dynamic diffuse optical tomography," Journal of Biomedical Optics, vol. 18, pp. 096012-096012, 2013.

[5] R. A. McLaughlin, B. C. Quirk, A. Curatolo, R. W. Kirk, L. Scolaro, D. Lorenser, et al., "Imaging of Breast Cancer With Optical Coherence Tomography Needle Probes: Feasibility and Initial Results," IEEE Journal of Selected Topics in Quantum Electronics, vol. 18, pp. 1184-1191, 2012.

[6] P.-L. Hsiung, D. R. Phatak, Y. Chen, A. D. Aguirre, J. G. Fujimoto, and J. L. Connolly, "Benign and Malignant Lesions in the Human Breast Depicted with Ultrahigh Resolution and Three-dimensional Optical Coherence Tomography," Radiology, vol. 244, pp. 865-874, 2007.

[7] B. F. Kennedy, R. A. McLaughlin, K. M. Kennedy, L. Chin, P. Wijesinghe, A. Curatolo, et al., "Investigation of Optical Coherence Microelastography as a Method to Visualize Cancers in Human Breast Tissue," Cancer Research, vol. 75, pp. 3236-3245, August 15, 2015 2015.

[8] D. Savastru, E. W. Chang, S. Miclos, M. B. Pitman, A. Patel, and N. Iftimia, "Detection of breast surgical margins with optical coherence tomography imaging: a concept evaluation study," Journal of Biomedical Optics, vol. 19, pp. 056001-056001, 2014.

[9] A. C. Sullivan, J. P. Hunt, and A. L. Oldenburg, "Fractal analysis for classification of breast carcinoma in optical coherence tomography," Journal of Biomedical Optics, vol. 16, pp. 066010-066010-6, 2011.

[10] A. M. Zysk and S. A. Boppart, "Computational methods for analysis of human breast tumor tissue in optical coherence tomography images," Journal of Biomedical Optics, vol. 11, pp. 054015-054015-7, 2006.

[11] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural Features for Image Classification," Systems, Man and Cybernetics, IEEE Transactions on, vol. SMC-3, pp. 610-621, 1973.

[12] M.-H. Horng, Y.-N. Sun, and X.-Z. Lin, "Texture feature coding method for classification of liver sonography," Computerized Medical Imaging and Graphics, vol. 26, pp. 33-42, 1/2/ 2002.

[13] G. van Soest, S. Koljenović, B. E. Bouma, G. J. Tearney, J. W. Oosterhuis, P. W. Serruys, et al., "Atherosclerotic tissue characterization in vivo by optical coherence tomography attenuation imaging," J Biomed Opt, vol. 15, pp. 011105-011105-9, 2010.

[14] Y. Gan, D. Tsay, S. B. Amir, C. C. Marboe, and C. P. Hendon, "Automated classification of optical coherence tomography images of human atrial tissue," Journal of Biomedical Optics, vol. 21, pp. 101407-101407, 2016.

[15] D. Barber, Bayesian Reasoning and Machine Learning: Cambridge University Press, 2012.

[16] X. Yao, Y. Gan, C. C. Marboe, and C. P. Hendon, "Myocardial Imaging Using Ultrahigh Resolution Spectral Domain Optical Coherence Tomography," Journal of Biomedical Optics, vol. 21, 2016.

[17] X. Yao, E. Chang, H. Hibshoosh, S. Feldman, and C. P. Hendon, "Towards in vivo high-resolution OCT based ductal imaging," in Biomedical Optics 2016, Fort Lauderdale, Florida, 2016, p. JTu3A.33.